



# De bé à bébé: le transfert d'apprentissage auditori-moteur pour interroger l'unité de production de la parole

Tiphaine Caudrelier, Pascal Perrier, Jean-Luc Schwartz, Christophe  
Savariaux, Amélie Rochet-Capellan

## ► To cite this version:

Tiphaine Caudrelier, Pascal Perrier, Jean-Luc Schwartz, Christophe Savariaux, Amélie Rochet-Capellan. De bé à bébé: le transfert d'apprentissage audiori-moteur pour interroger l'unité de production de la parole. JEP-TALN-RECITAL 2016 - conférence conjointe 31e Journées d'Études sur la Parole, 23e Traitement Automatique des Langues Naturelles, 18e Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues, AFCEP, Jul 2016, Paris, France. hal-01343390

**HAL Id: hal-01343390**

**<https://hal.science/hal-01343390>**

Submitted on 8 Jul 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# De bé à bébé : le transfert d'apprentissage audiotri-moteur pour interroger l'unité de production de la parole

Tiphaine Caudrelier<sup>1</sup> Pascal Perrier<sup>1</sup> Jean-Luc Schwartz<sup>1</sup> Christophe Savariaux<sup>1</sup> Amélie  
Rochet-Capellan<sup>1</sup>

(1) Gipsa-Lab, 11 rue des Mathématiques, 38400 Saint Martin d'Hères, France

Tiphaine.caudrelier@gipsa-lab.fr, Amélie.Rochet-Capellan@gipsa-lab.fr

## RESUME

---

La parole est souvent décrite comme une mise en séquence d'unités associant des représentations linguistiques, sensorielles et motrices. Le lien entre ces représentations se fait-il de manière privilégiée sur une unité spécifique ? Par exemple, est-ce la syllabe ou le mot ? Dans cette étude, nous voulons contraster ces deux hypothèses. Pour cela, nous avons modifié chez des locuteurs du français la production de la syllabe « bé », selon un paradigme d'adaptation audiotri-motrice, consistant à perturber le retour auditif. Nous avons étudié comment cette modification se transfère ensuite à la production du mot « bébé ». Les résultats suggèrent un lien entre représentations linguistiques et motrices à plusieurs niveaux, à la fois celui du mot et de la syllabe. Ils montrent également une influence de la position de la syllabe dans le mot sur le transfert, qui soulève de nouvelles questions sur le contrôle sériel de la parole.

## ABSTRACT

---

**From sensorimotor experience to speech unit.**

Speech is often described as a sequence of units associating linguistic, sensory and motor representations. Are these representations linked at the level of a specific unit, for example, the syllable or the word? In the present study, we contrast these two hypotheses. We modified the production of the syllable “bé” (/be/) in French speakers using an auditory-motor adaptation paradigm that consists in altering the speakers' auditory feedback. We studied how this modification then transfers to the production of the word “bébé” (/bebe/). The results suggest a link between linguistic and motor representations both at the word and the syllable level. They also show an effect of the position of the syllable in the transfer word, which raises new interrogations about serial control of speech.

---

**MOTS-CLES :** Perturbations, adaptation audiotri-motrice, transfert, unité de parole

**KEYWORDS:** Perturbations, auditory-motor adaptation, transfer, speech unit

---

## 1 Introduction

La parole peut être décrite comme une mise en séquence d'unités linguistiques hiérarchisées entre elles, telles que les phrases, les mots, les syllabes ou les phonèmes. Ces unités seraient également de

nature sensorimotrice, c'est-à-dire associant des représentations sensorielles, notamment auditives, et motrices. Au niveau du système moteur, elles correspondent à des gestes articulatoires, qui se succèdent et parfois se recouvrent. Mais quelle est la nature du lien entre ces représentations linguistiques et motrices ? A quel niveau s'effectue la correspondance entre les deux ?

Différentes perspectives existent sur cette question. La syllabe Consonne-Voyelle (CV) a souvent été présentée comme l'unité principale de production de la parole. Par exemple la Théorie « Frame then Content » (Cadre puis Contenu) de MacNeilage (1998) suggère que la syllabe CV trouve ses racines dans les oscillations mandibulaires (notamment présentes dans la mastication), et pourrait ainsi être à l'origine de notre langage articulé. Par ailleurs, en psycholinguistique, Crompton (1982) a suggéré l'existence d'un répertoire mental, appelé syllabaire, qui contiendrait les gestes articulatoires nécessaires à la production de chaque syllabe. Levelt (1999) a développé cette idée dans son modèle de production de la parole où la syllabe serait le niveau auquel le message est décomposé afin d'être traduit, à l'aide du syllabaire, en commandes motrices. Il s'appuie sur l'observation des lapsus phonologiques (inversion d'unités dans une phrase), et sur des effets de fréquence des syllabes sur la vitesse de prononciation des mots.

Néanmoins, l'hypothèse que la syllabe serait l'unité majeure de production de la parole est contestée. En particulier, la grande variabilité observée dans la production d'une syllabe par un locuteur en fonction du contexte de communication demande explication. La théorie de l'exemplaire (Medin et Schaffer 1978) propose une vision alternative permettant de prendre en compte ces variabilités. Elle a d'abord été développée comme un modèle général de perception et de catégorisation, et repose sur le principe que chaque élément perçu d'une catégorie est mémorisé, plutôt que le seul prototype de la catégorie. Ainsi, un stimulus donné serait comparé à tous les éléments en mémoire pour être catégorisé. Cette théorie a été ensuite étendue à l'identification et la reconnaissance de mots par Goldinger (1998) puis à la production de parole (Bybee 2002). Elle est également au cœur des propositions de Vihman et Croft (2007) sur le développement de la phonologie chez l'enfant, dans une perspective où la syllabe émerge de la production des premiers mots.

Notre objectif est d'interroger la nature de l'unité de production de la parole. Pour évaluer quelle unité assure la correspondance entre une séquence de parole et les gestes articulatoires la produisant, une possibilité est de modifier cette correspondance pour une instance donnée, et de voir comment cette modification s'étend à d'autres instances. Par exemple, si la syllabe est l'unité de production assurant le lien avec les commandes motrices, alors si l'on modifie la production d'une syllabe dans un contexte donné, cette modification devrait se généraliser à la production de cette syllabe dans n'importe quel mot. A l'inverse, si le mot est la plus petite unité représentée, alors il ne devrait pas y avoir de transfert à la même syllabe dans un autre mot. Ce sont ces deux hypothèses que nous allons tester, grâce à un paradigme d'adaptation et de transfert de l'apprentissage sensorimoteur.

L'idée de modifier, à l'aide d'une perturbation, la relation entre les unités perceptives et motrices vient de la recherche sur le contrôle moteur, notamment le contrôle du bras (Mattar et Ostry 2007). Elle a été adaptée à la production de la parole par Houde et Jordan (1998). Ils ont développé le paradigme d'adaptation auditori-motrice, par analogie avec l'étude de la recalibration visuo-motrice consécutive au port de lunettes prismatiques. Le paradigme est le suivant. Les participants doivent répéter un mot de type Consonne-Voyelle-Consonne (CVC) tandis qu'ils entendent leur voix via un casque. Leur retour auditif est modifié en temps réel. L'altération consiste à décaler le formant f1 (et/ou f2) de manière à transformer la voyelle en une autre (par exemple, « head » devient « had »). Les participants modifient alors leur prononciation, en compensant partiellement la perturbation, se rapprochant ainsi de « hid ». Une phase d'entraînement, constituée de nombreuses répétitions,

aboutit à des recalibrations auditori-motrices, ou apprentissage, qui perdure après l'arrêt de la perturbation. Cet apprentissage peut être évalué en comparant les formants produits avant et après l'entraînement. C'est ce qu'on appelle le post-effet. Houde et Jordan (2002) ont aussi montré que l'apprentissage réalisé sur une certaine instance se transfère partiellement à la même voyelle dans d'autres mots, voire à des voyelles différentes. Le transfert à la même voyelle dans d'autres contextes serait selon eux le signe de l'existence d'une représentation mentale du phonème.

Ce paradigme d'adaptation et de transfert de l'apprentissage auditori-moteur a été repris dans d'autres études, avec des mots ou pseudo-mots de type CVC. Par exemple, dans Rochet-Capellan et al. (2012), chaque groupe de participants est entraîné sur un mot de la forme Consonne-Voyelle-/n/ différent (par exemple « pan » et « ten »), puis le transfert au mot « pen » est évalué. Cette étude montre qu'un transfert significatif intervient pour la plupart des groupes mais que son amplitude dépend du mot d'entraînement. Une généralisation large mais de faible amplitude a été aussi observée en Mandarin, à partir d'une adaptation sur la triphongue /iau/ (Cai et al., 2010). Purcell et Munhall (2006) ont estimé l'influence de différents paramètres du protocole expérimental (amplitude de la perturbation, durée de l'entraînement) sur l'adaptation et sur la vitesse de disparition des effets de l'apprentissage après arrêt de la perturbation (ou durée du post-effet). D'autres recherches ont contribué à comprendre comment l'adaptation auditori-motrice est influencée par des facteurs linguistiques, tels que le statut lexical, i.e. mot ou pseudo-mot, (Bourguignon et al., 2014) ou la fréquence des mots (Frank 2011). Cependant, la plupart de ces études ont été réalisées avec des locuteurs de l'anglais. A notre connaissance, l'utilisation de ce paradigme expérimental pour contraster l'importance du mot et de la syllabe CV en production de parole est nouvelle.

## **2 Méthode**

### **2.1 Participants, équipement et tâche**

36 participants âgés de 18 à 35 ans, de langue maternelle française, ont pris part à l'expérience. Ils n'avaient pas de trouble de l'audition, ni de trouble du langage, et étaient naïfs quant au but de l'expérience. Ils ont reçu une carte cadeau de 15 euros pour leur participation. Les locuteurs étaient assis devant un écran, dans une chambre sourde. Ils portaient un casque équipé d'un microphone (Sennheiser HME 26-II-600). Des mots apparaissaient à l'écran pour une durée d'1.1s. Ils avaient pour consigne de lire les mots à haute voix, de manière naturelle, sans murmurer ni crier.

### **2.2 Design expérimental et stimuli**

L'objectif était de comparer le transfert de l'apprentissage auditori-moteur au niveau du mot, et de la syllabe CV. Pour contraster ces 2 niveaux, nous avons choisi d'utiliser pour mot d'entraînement une syllabe CV, et d'étudier d'une part le post-effet pour le même mot monosyllabique, et d'autre part le transfert à un mot de la forme CVCV répétant cette syllabe. Le mot d'entraînement sélectionné était « bé » pour tous les participants. A l'arrêt de la perturbation auditive, un groupe de sujets a prononcé « bé », et l'autre « bébé ». Notre hypothèse est que si la syllabe est l'unité de production principale, on devrait observer un transfert aux deux voyelles de « bébé » de même amplitude que le post-effet observé sur « bé ». A l'inverse, si le mot est l'unité principale, il ne devrait y avoir aucun transfert à « bébé ». Le choix de la syllabe /be/ devait satisfaire à plusieurs critères. Le choix de la voyelle /e/ était contraint par des aspects techniques liés à la perturbation auditive appliquée. La voyelle choisie

devait nécessairement être située entre deux autres voyelles dans l’espace formantique. La détection des formants étant difficile sur les voyelles arrière /o/ et /ɔ/, le choix s’est restreint aux voyelles avant /e/ et /ɛ/. Par ailleurs, la consonne « b » induit une faible coarticulation. En français, la syllabe « bé » est aussi fréquente en première position d’un mot qu’en milieu ou fin de mot. Enfin, la fréquence du mot d’entraînement et de ses voisins auditifs et articulatoires (« bi » et « baie ») peut influencer l’adaptation auditori-motrice, d’après Frank (2011). Nous avons donc cherché à minimiser la fréquence de ces mots.

### 2.3 Transformation du retour auditif

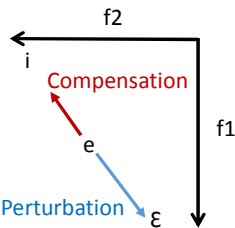


FIGURE 1: La perturbation auditive consiste en un décalage des formants f1 et f2 de manière à transformer le /e/ en /ɛ/. On s’attend à ce que les locuteurs compensent la perturbation en prononçant un /e/ tendant vers le /i/.

Pendant toute l’expérience, les participants parlaient dans un micro et entendaient leur voix dans le casque avec un volume d’environ 70dB, mixée avec un bruit blanc à environ 50dB. Leur voix était enregistrée en parallèle avec un taux d’échantillonnage de 48kHz. Une perturbation en temps réel des formants f1 et f2 était réalisée grâce au logiciel Audapter (Cai et al., 2008). La perturbation a été paramétrée de manière à transformer la voyelle /e/ en /ɛ/. Elle consistait à augmenter f1 de 27% et à diminuer f2 de 10% dans la voyelle /e/. La perturbation créait un retard de 14ms au niveau du retour dans le casque, non perceptible par les participants.

### 2.4 Procédure expérimentale

Un pré-test permettait d’explorer le triangle vocalique des participants, en leur faisant prononcer des syllabes contenant les voyelles /a/, /ɛ/, /e/, /i/, /u/ et /y/ précédées de la consonne /b/. Le pré-test contenait aussi des paires minimales ayant pour but d’évaluer le contraste /e/-/ɛ/ en production chez les participants (par exemple « épée » vs « épais ») (blocs 1, 2, 3, Fig. 2).

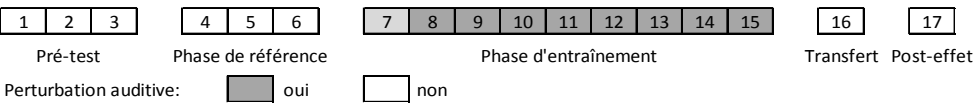


FIGURE 2: L’expérience était constituée d’un pré-test et de 4 phases expérimentales. Une perturbation auditive (en gris) était appliquée pendant la phase d’entraînement.

L’expérience se déroulait par blocs de 20 essais entrecoupés par des pauses. Dans une phase dite de « référence », les participants répétaient 20 fois le mot « bé » (bloc 4, Fig. 2) puis 20 fois le mot « bébé », quel que soit leur groupe (bloc 5). Ensuite, la phase d’entraînement était constituée d’un

bloc de rampe pendant lequel la perturbation était instaurée graduellement jusqu'à son amplitude maximale (bloc 7), suivi de 8 blocs avec la perturbation complète telle que décrite précédemment (blocs 8-15). Tous les participants prononçaient « bé ». A l'issue de l'entraînement, la perturbation était arrêtée, et les participants entendaient à nouveau leur voix non modifiée. Pendant un bloc, dit « de transfert » (16), un groupe continuait à dire « bé » tandis que l'autre groupe prononçait « bébé ». Enfin, dans le bloc de post-effet (17), les deux groupes prononçaient « bé » afin d'évaluer l'adaptation résiduelle après la phase de transfert.

## 2.5 Analyse des données

Pour chaque voyelle produite, les formants f1 et f2 ont été estimés en prenant une fenêtre de 30 ms au milieu de la partie stable de la voyelle. Pour évaluer l'évolution au cours de l'expérience, les formants ont été exprimés en pourcentage de changement par rapport à leur valeur moyenne pour la même voyelle dans le même mot et pour le même sujet avant la perturbation. Par exemple, le f1 du 2<sup>ème</sup> /e/ de « bébé » a été rapporté à la valeur moyenne de f1 du 2<sup>ème</sup> /e/ de « bébé » dans la phase de référence. Comme la perturbation affectait à la fois f1 et f2, la mesure principale choisie pour évaluer le transfert est f2-f1, avec f1 et f2 exprimés en bark ( $f_{bark} = 7 * \text{argsinh}(f\text{Hz}/650)$ ).

Si la syllabe est l'unique unité de production de la parole (H1), on ne devrait observer aucun effet du mot de test (« bé » ou « bébé ») ni de la position de la syllabe dans « bébé » sur l'amplitude du changement de f2-f1 observé dans le bloc de transfert. A l'inverse, selon l'hypothèse de spécificité du mot (H2), on devrait observer un post-effet sur « bé » mais pas de transfert à « bébé ». Pour estimer l'adaptation de chaque participant, la valeur de f2-f1 des 40 derniers essais de la phase d'entraînement a été comparée aux valeurs de f2-f1 dans la phase de référence en utilisant un test de Student apparié unilatéral. Le transfert a été estimé en faisant un test de Student sur la valeur du changement sur f2-f1 par rapport à la phase de référence, en contrastant d'une part le groupe contrôle (prononçant « bé ») avec la 1<sup>ère</sup> voyelle de « bébé » (test de Student non apparié unilatéral) et d'autre part la 1<sup>ère</sup> avec la 2<sup>ème</sup> voyelle de « bébé » (test de Student apparié bilatéral).

## 3 Résultats

### 3.1 Adaptation et sélection des participants

Nous nous attendions à ce que les participants compensent la perturbation du retour auditif en augmentant f2-f1 dans leurs productions de « bé » (aboutissant à une syllabe plus proche de /bi/). Comme cette étude se focalise sur le transfert, seuls les 27 participants présentant une adaptation significative ont été inclus dans les analyses, soit soit un groupe de 13 participants (3 femmes) pour le groupe contrôle (transfert 'bé') et un groupe de 14 participants (3 femmes) pour le groupe test (transfert 'bébé').

L'évolution moyenne de f2-f1 au cours de l'expérience est représentée sur la Figure 3. En réponse à la diminution de f2-f1 dans le retour auditif, on observe une augmentation progressive de f2-f1 en production qui se stabilise en fin d'apprentissage. L'adaptation a été évaluée en comparant les 40 derniers essais de la phase d'entraînement avec les 20 essais de la phase de référence. En moyenne, on observe une augmentation de f2-f1 de 7.4% (erreur standard :  $\pm 0.3\%$ ). Celle-ci ne dépend pas du groupe. Elle correspond à une compensation de la perturbation d'environ 25%. Ainsi, l'adaptation est significative et homogène entre les deux groupes. L'amplitude de la compensation est comparable à celle qui a été observée dans des travaux antérieurs (Rochet-Capellan et al., 2012).

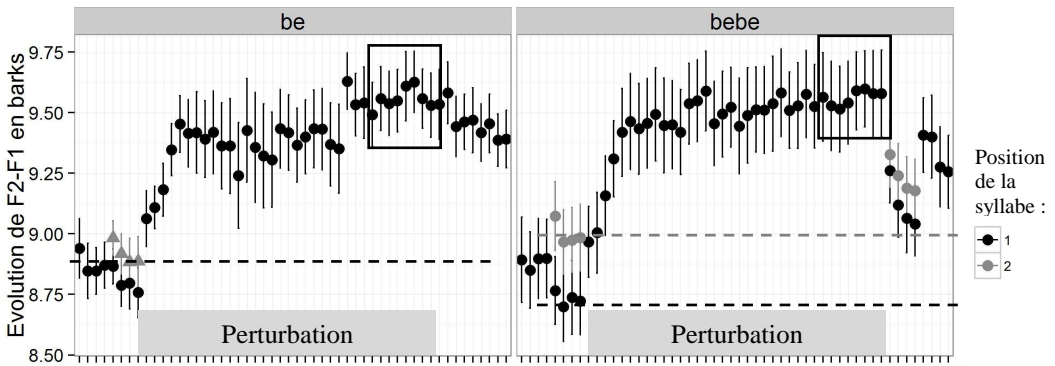


FIGURE 3: Evolution de  $f_2-f_1$  en barks au cours de l'expérience, pour les sujets adaptés de chaque groupe (à gauche 'bé', à droite 'bébé'). Chaque point représente la moyenne du groupe pour 5 essais avec l'intervalle de confiance. Les lignes pointillées donnent les valeurs de la phase référence pour 'bé' et 'bébé'. Les rectangles montrent les essais considérés pour quantifier l'adaptation.

### 3.2 Post-effet sur « bé » vs transfert sur « bébé »

Le transfert a été mesuré en comparant les 20 essais qui suivent l'arrêt de la perturbation du retour auditif (bloc 16 de la Fig. 2) avec les valeurs de  $f_2-f_1$  de la phase de référence, pour chacun des mots et des syllabes. Les résultats sont représentés pour chaque groupe, mot et syllabe, sur la Figure 4. L'amplitude du changement de  $f_2-f_1$  sur « bé » dans le groupe contrôle est de 7.0% ( $\pm 1.0\%$ ). Pour le groupe ayant prononcé « bébé », le changement par rapport à la phase de référence est respectivement de 4.6% ( $\pm 0.7\%$ ) pour la 1<sup>ère</sup> syllabe et 2.6% ( $\pm 0.7\%$ ) pour la 2<sup>ème</sup> syllabe. Un t-test révèle que le post-effet sur « bé » est significativement plus élevé que le transfert à la 1<sup>ère</sup> syllabe de « bébé » ( $t(25) = 1.97$ ,  $p = 0.03$ ). De plus, le transfert à la 1<sup>ère</sup> syllabe de « bébé » est significativement plus grand que le transfert à la 2<sup>ème</sup> syllabe ( $t(13) = 3.25$ ,  $p = 0.006$ ).

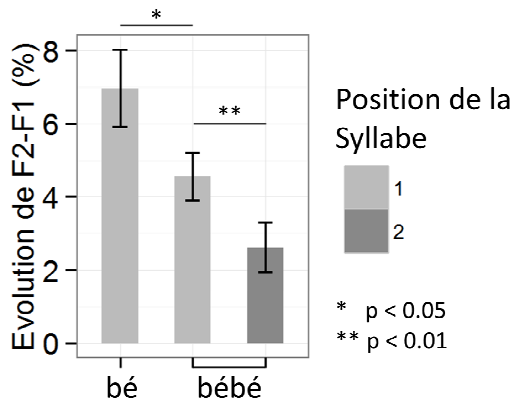


FIGURE 4:  $f_2-f_1$  moyen dans le bloc de transfert (16), en pourcentage de changement par rapport aux valeurs de référence, par groupe (« bé » ou « bébé ») et par syllabe (1<sup>ère</sup> ou 2<sup>ème</sup> position).

Dans le bloc de post-effet (bloc 17), l'amplitude du changement de  $f_2$ - $f_1$  sur « bé » par rapport à la phase de référence a été mesurée pour les deux groupes. Chez le groupe qui a prononcé « bé » dans le bloc précédent, le post-effet est de 6.1% ( $\pm 1.0\%$ ). Il est de 5.0% ( $\pm 0.6\%$ ) chez le groupe ayant prononcé « bébé ». Il n'y a pas de différence significative entre les deux groupes ( $p=0.27$ ).

## 4 Discussion

Avant de discuter des résultats sur le transfert, il est important de préciser un certain nombre d'hypothèses sur la nature de l'adaptation sensorimotrice. D'abord, on peut se demander si l'adaptation est de nature sensorimotrice ou bien uniquement perceptive. Si l'adaptation était uniquement perceptive, alors le locuteur s'habituerait à un /e/ qui ressemble à un /E/ (c'est le sens de la perturbation auditive). On pourrait donc s'attendre à ce que lorsque l'on arrête la perturbation, le participant se mette à prononcer une syllabe proche de /bE/, pour rester dans la même catégorie perceptive. Or c'est l'inverse qui se produit. Le locuteur prononce une syllabe proche de /bi/, qui est éloignée d'un point de vue auditif mais identique d'un point de vue moteur à ce qu'il a appris pendant l'entraînement. Il ne s'agit donc pas d'une adaptation des catégories auditives.

D'autre part, on peut également se demander si l'adaptation résulte d'une recalibration automatique chez les participants, ou bien si ces derniers adoptent une stratégie de compensation de la perturbation, en supposant ce que l'expérimentateur attend d'eux. Les participants ont répondu à un questionnaire à la fin de l'expérience. On leur demande notamment s'ils ont perçu quelque chose de bizarre au niveau de leur retour auditif, s'ils ont eu l'impression que leur voix avait été modifiée, et si oui comment. Sur 36 participants, 21 ont déclaré avoir remarqué une modification de leur voix, et seulement 10 participants (soit 28%) ont mentionné une transformation de /e/ vers /E/. Selon 5 d'entre eux, cette transformation était occasionnelle (par exemple au début de certaines séquences). Ainsi la plupart des participants n'avaient aucune idée de ce que l'expérimentateur pouvait attendre d'eux, puisqu'ils n'avaient même pas identifié la perturbation. Par ailleurs une étude a montré que les participants s'adaptent à une perturbation auditive même quand on leur demande de ne pas s'adapter (Munhall et al., 2009).

### 4.1 Que nous apprend le transfert sur l'unité de production de la parole?

Notre objectif était de contraster les rôles respectifs du mot et de la syllabe en production de parole, en étudiant le transfert de l'apprentissage sensorimoteur. Les résultats représentent un mélange de nos deux hypothèses. Le transfert significatif aux voyelles de « bébé » plaide pour l'existence d'une représentation d'une unité de production de la parole plus petite que le mot, qui pourrait être la syllabe CV. Néanmoins, la différence d'amplitude entre le post-effet sur « bé » et le transfert à « bébé » questionne l'existence d'un syllabaire mental qui représenterait l'unique lien entre une séquence de parole cible et les gestes articulatoires permettant de la produire. Ainsi, le lien se ferait à de multiples niveaux, incluant à la fois le mot et la syllabe, et probablement d'autres unités.

Il est à noter que les sujets du groupe « bébé » ont prononcé la syllabe « bé » deux fois plus que les autres dans le bloc de transfert. On pourrait donc s'attendre à ce que les effets de l'entraînement disparaissent plus rapidement dans ce groupe. Néanmoins cet effet possible de la répétition n'est pas à lui-seul à l'origine de l'effet observé entre le post-effet sur « bé » et le transfert à « bébé ». En effet, le post-effet mesuré dans le bloc suivant est équivalent dans les deux groupes.



## 4.2 Effet de la position de la syllabe

On observe dans cette étude une influence de la position de la syllabe dans le mot sur le transfert de l'apprentissage sensorimoteur. Cette différence va de nouveau à l'encontre de la conception d'un syllabaire contenant une représentation unique, un prototype, de la syllabe et des gestes articulatoires qui lui sont associés. En effet, le transfert est significativement plus élevé sur la 1<sup>ère</sup> syllabe que sur la 2<sup>ème</sup> syllabe. Nous proposons plusieurs explications à cette observation.

D'abord, cet effet de position pourrait être lié à une influence de la prosodie sur le transfert de l'apprentissage sensorimoteur. Une syllabe plus accentuée pourrait être produite de manière plus précise. Nous avons observé chez les participants de cette étude que la première syllabe de bébé est significativement plus accentuée que la 2<sup>ème</sup> selon deux indices prosodiques : la fréquence fondamentale  $f_0$  et l'énergie acoustique sont plus élevées sur la 1<sup>ère</sup> syllabe que sur la 2<sup>ème</sup>. À l'inverse, la durée de la 2<sup>ème</sup> voyelle est plus longue que la 1<sup>ère</sup>. Ces observations sont valables à la fois dans la phase de phase de référence et dans le bloc de transfert.

L'effet de position observé pourrait dépendre aussi de la perturbation. La phase d'entraînement a été réalisée avec un mot monosyllabique. La syllabe perturbée était donc la première (et l'unique). Dans ce contexte, le locuteur pourrait peut-être apprendre à corriger particulièrement la 1<sup>ère</sup> syllabe d'un mot ou d'une séquence. Aucune étude utilisant le paradigme d'adaptation auditori-motrice n'a été réalisée à notre connaissance chez l'humain avec des mots de plusieurs syllabes. Cependant, une étude du transfert de l'apprentissage auditori-moteur dans une séquence de syllabes a été réalisée chez les oiseaux. Hoffman et Sober (2014) ont utilisé le paradigme d'adaptation auditori-motrice sur des vocalisations de diamants mandarins. Ceux-ci étaient équipés d'un mini casque leur fournissant un retour auditif. Une perturbation de leur fréquence fondamentale était appliquée sur une syllabe dans une position précise de leur vocalisation. Cette perturbation a induit une adaptation allant dans le sens d'une compensation, comme chez l'humain. Après l'arrêt de la perturbation, un transfert a été observé pour les syllabes identiques à la syllabe sur laquelle la perturbation avait été appliquée, quelle que soit leur position dans la séquence. Le transfert s'étendait partiellement aux syllabes voisines dans la séquence de test. Chez l'humain, l'utilisation de perturbations sélectives en termes de position dans le mot ouvrirait de nouvelles perspectives intéressantes sur l'étude des programmes moteurs nécessaires à la réalisation de séquences de parole.

## Remerciements

Ces recherches ont bénéficié du soutien financier du Conseil Européen de la Recherche sous le septième programme-cadre de l'Union Européenne (FP7/2007-2013 Grant Agreement no. 339152, "Speech Unit(e)s", PI: Jean-Luc-Schwartz).

## Références

- BYBEE, J. (2002). Phonological Evidence for Exemplar Storage of Multiword Sequences. *Studies in Second Language Acquisition*, 24.
- CAI, S., BOUCEK, M., GHOSH, S. S., GUENTHER, F. H., & PERKELL, J. S. (2008). A System for Online Dynamic Perturbation of Formant Trajectories and Results from Perturbations of the Mandarin. *International Seminar on Speech Production 2008*, 65–68.

- CAI, S., GHOSH, S. S., GUENTHER, F. H., & PERKELL, J. S. (2010). Adaptive auditory feedback control of the production of formant trajectories in the Mandarin triphthong /iau/ and its pattern of generalization. *The Journal of the Acoustical Society of America*, 128, 2033–2048.
- FRANK, A. F. (2011). Integrating Linguistic, Motor, and Perceptual Information in Language Production. *Dissertation Abstracts International, B: Sciences and Engineering*, 72, 2454.
- GOLDINGER, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279.
- HOFFMANN, L. A., & SOBER, S. J. (2014). Vocal Generalization Depends on Gesture Identity and Sequence. *Journal of Neuroscience*, 34(16), 5564–5574.
- HOUE, J. (2002). Sensorimotor Adaptation of Speech I : Compensation and Adaptation. *JSLHR*, 45(April 2002).
- HOUE, J. F., & JORDAN, M. I. (1998). Sensorimotor adaptation in speech production. *Science (New York, N.Y.)*, 279(1998), 1213–1216.
- LEVELT, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223–232.
- MACNEILAGE, P. F. (1998). The frame/content theory of evolution of speech production. *The Behavioral and Brain Sciences*, 21(4), 499–511; discussion 511–546.
- MATTAR, A. A G., & OSTRY, D. J. (2007). Modifiability of generalization in dynamics learning. *Journal of Neurophysiology*, 98(6), 3321–3329.
- MEDIN, D. L., & SCHAFER, M. M. (1978). Context theory of classification learning. *Psychological review*, 85(3), 207.
- MUNHALL, K. G., MACDONALD, E. N., BYRNE, S. K., & JOHNSRUDE, I. (2009). Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *The Journal of the Acoustical Society of America*, 125(1), 384–390.
- PURCELL, D. W., & MUNHALL, K. G. (2006). Adaptive control of vowel formant frequency: evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America*, 120, 966–977.
- ROCHET-CAPELLAN, A., RICHER, L., & OSTRY, D. J. (2012). Nonhomogeneous transfer reveals specificity in speech motor learning. *Journal of Neurophysiology*, 107, 1711–1717.
- VIHMAN, M. M., & CROFT, W. (2007). Phonological development: Toward a “radical” templatic phonology. *Linguistics*, 45, 683–725.